Chongyang Shi

Gainesville, Florida | c.shi@ufl.edu | 352 226 9039 | Homepage | Google Scholar | LinkedIn | GitHub

Summary

Ph.D. researcher specializing in **LLM**, **Agentic Reinforcement Learning**, and **Multi-Agent Systems**. My work focuses on user and environment modeling, active information acquisition, tool-using and reasoning, complex and multi-step tasks, with applications to AI systems and foundation models. Experienced in PyTorch-based GPU training and published at top AI conferences including IJCAI, AAMAS, ICCPS, CDC, and IEEE L-CSS.

Education

University of Florida, Ph.D. in Electrical and Computer Engineering — Aug 2022 – May 2027

Southern University of Science and Technology, B.S. in Statistics — Sept 2018 – June 2022

Experience

Pinterest Research Labs, Machine Learning Engineer Intern — May 2026 – Aug 2026

Project: Scalable long-sequence modeling for large-scale retrieval systems

- Developed long-sequence user modeling techniques for **ads retrieval** systems using **MMOE** and **transformer**.
- Designed and evaluated different **attention architectures**, and **context-aware** filter methods to solve the **memory bottleneck**. Trained the model on **AWS** cloud service using **distributed** method.
- Changed the **time-based positional encoding** to incorporate the long sequence features.
- Improved the recall for low relevance bucket by **1.28%** compared to production model.

Selected Projects

Co-Evolving LLM Teachers and Robot Reinforcement Learning — Sep 2026 – Present

- Developed a **co-evolving LLM-robot RL framework** for adaptive curriculum generation and robot learning.
- Built a **closed-loop training pipeline** where robot feedback guides LLM curriculum refinement.
- Trained the **LLM teacher** with policy gradient using automata-based verifiable rewards based on robot success.
- Generated **language-aligned reward machines** with subgoals and policy-invariant reward shaping.
- Designed benchmarks in **XLand-MiniGrid, Craftax, ZoneEnv, and Isaac Lab**.

Privacy-Preserving LLM Agents via Information-Theoretic Post-Training, — Aug 2026 – Present

- Developed an information-theoretic framework for hiding **sensitive attributes** in **long-horizon** shop agents.
- Built **tool-using** agents with **Qwen3-0.6B** and trained attackers to detect privacy leakage from agent trajectories.
- Applied **SFT + GRPO** to optimize task performance while minimizing trajectory-level information leakage.
- Evaluated privacy–utility tradeoffs using attacker accuracy, conditional entropy, and agent task success.
- Reduced **information leakage** by 20% compared to baselines while satisfying performance requirements.

Multi-Agent Information Acquisition and User Intent Inference, — June 2025 – Oct 2025

- Developed a multi-agent framework for User Intent inference from distributed observations.
- Proposed an information-theoretic method for **agent selection** and **coordinated information acquisition**.
- Studied collaboration strategies among **fully decentralized** agents for efficient information gathering.
- Achieved 88% inference accuracy while reducing computation by 5$\times$ compared to **PPO-like** baselines.

Active Inference for User Intent Identification, — Jan 2025 – May 2025

- Developed an **active inference** framework for personalized LLM agents with latent user types.
- Designed information-directed (mutual information) rollout selection to identify the most informative interactions for user intent and preference inference.
- Optimized **information gain** over user types and downstream task performance using the policy gradient.
- Improved user-type inference accuracy from **79.1% to 99.6%** over reward-only policies.

Relevant Skills

Deep Learning: Representation learning, multimodal encoders, variational autoencoders, generative models, prompt optimization, synthetic data generation, post-training, reasoning models, LLM evaluation
Tools: PyTorch, GPU training, Autonomous agents, planning, tool use, memory, reasoning
Programming: Python, Java, C/C++, CUDA (basic), MATLAB
Courses: Mathematical statistics, Bayesian statistics, information theory, time series analysis, stochastic processes, machine learning, deep learning

Publications

Journal Papers

Active Perception With Initial-State Uncertainty: A Policy Gradient Method
Chongyang Shi, Shuo Han, Michael R. Dorothy, Jie Fu
IEEE Control Systems Letters (L-CSS), PDF

Conference Papers

Quantitative Planning with Action Deception in Concurrent Stochastic Games
Chongyang Shi, Shuo Han, Jie Fu
International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2023), PDF

Synthesis of Opacity-Enforcing Winning Strategies Against Colluded Opponent
Chongyang Shi, Abhishek N. Kulkarni, Hazhar Rahmani, Jie Fu
IEEE Conference on Decision and Control (CDC 2023), PDF

Information-Theoretic Opacity-Enforcement in Markov Decision Processes
Chongyang Shi, Yuheng Bu, Jie Fu
International Joint Conference on Artificial Intelligence (IJCAI 2024), PDF

Covert Planning against Imperfect Observers
Haoxiang Ma, **Chongyang Shi**, Shuo Han, Michael R. Dorothy, Jie Fu
International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2024), PDF

Synthesis of Dynamic Masks for Information-Theoretic Opacity in Stochastic Systems
Sumukha Udupa, **Chongyang Shi**, Jie Fu
ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS 2025), PDF

Active Inference through Incentive Design in Markov Decision Processes
Xinyi Wei, **Chongyang Shi**, Shuo Han, Ahmed H. Hemida, Charles A. Kamhoua, Jie Fu,
International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2026), PDF

IMAS^2: Joint Agent Selection and Information-Theoretic Coordinated Perception In Dec-POMDPs
Chongyang Shi, Wesley A. Suttle, Michael Dorothy, Jie Fu,
International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2026), PDF

Preprints

Integrated Control and Active Perception in POMDPs for Temporal Logic Tasks and Information Acquisition
Chongyang Shi, Michael R. Dorothy, Jie Fu, arXiv, (submitted to IROS 2026)

Policy Gradient Methods for Information-Theoretic Opacity in Markov Decision Processes
Chongyang Shi, Sumukha Udupa, Michael R. Dorothy, Shuo Han, Jie Fu, arXiv, (submitted to IEEE TAC)

C-IDS: Solving Contextual POMDP via Information-Directed Objective
Chongyang Shi, Michael R. Dorothy, Jie Fu, arXiv, (submitted to NeurIPS 2026)

When to Stop: Tuning-Free, Transferable Depth Control for Planning on Frozen World Models
Tong Zhang, **Chongyang Shi**, Yun Peng, Tao Xie, arXiv, (submitted to AAAI 2027)